



Inferring the Dysconnection Syndrome in Schizophrenia: Interpretational Considerations on Methods for the Network Analyses of fMRI Data

Brian H. Silverstein¹, Steven L. Bressler² and Vaibhav A. Diwadkar^{1*}

¹ Department of Psychiatry and Behavioral Neurosciences, Brain Imaging Research Division, Wayne State University, Detroit, MI, USA, ² Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA

OPEN ACCESS

Edited by:

Stefan Borgwardt,
University of Basel, Switzerland

Reviewed by:

Karl Friston,
University College London, UK
Sophia Frangou,
Icahn School of Medicine
at Mount Sinai, USA

*Correspondence:

Vaibhav A. Diwadkar
vdiwadka@med.wayne.edu

Specialty section:

This article was submitted to
Neuroimaging and Stimulation,
a section of the journal
Frontiers in Psychiatry

Received: 16 June 2016

Accepted: 15 July 2016

Published: 03 August 2016

Citation:

Silverstein BH, Bressler SL and
Diwadkar VA (2016) Inferring the
Dysconnection Syndrome in
Schizophrenia: Interpretational
Considerations on Methods for the
Network Analyses of fMRI Data.
Front. Psychiatry 7:132.
doi: 10.3389/fpsy.2016.00132

Schizophrenia has long been considered one of the most intractable psychiatric conditions. Its etiology is likely polygenic, and its symptoms are hypothesized to result from complex aberrations in network-level neuronal activity. While easily identifiable by psychiatrists based on clear behavioral signs, the biological substrate of the disease remains poorly understood. Here, we discuss current trends and key concepts in the theoretical framework surrounding schizophrenia and critically discuss network approaches applied to neuroimaging data that can illuminate the correlates of the illness. We first consider a theoretical framework encompassing basic principles of brain function ranging from neural units toward perspectives of network function. Next, we outline the strengths and limitations of several fMRI-based analytic methodologies for assessing *in vivo* brain network function, including undirected and directed functional connectivity and effective connectivity. The underlying assumptions of each approach for modeling fMRI data are treated in some quantitative detail, allowing for assessment of the utility of each for generating inferences about brain networks relevant to schizophrenia. fMRI and the analyses of fMRI signals provide a limited, yet vibrant platform from which to test specific hypotheses about brain network dysfunction in schizophrenia. Carefully considered and applied connectivity measures have the power to illuminate loss or change of function at the network level, thus providing insight into the underlying neurobiology which gives rise to the emergent symptoms seen in the altered cognition and behavior of schizophrenia patients.

Keywords: brain networks, fMRI methods, schizophrenia, connectivity analysis, dysconnection syndrome

INTRODUCTION

Schizophrenia is the consummate “epigenetic puzzle” (1). Psychiatrists, for the most part, know it when they see it in the clinic (2), yet its biological *origins* are utterly obscure, given that its etiology and genetic bases are poorly understood. We seem to *know* much but *understand* very little (3). Scientists have settled on the view that schizophrenia is a polygenic disorder (4, 5) where multiple genes (that themselves exert pleiotropic effects) confer vulnerability, but wherein

the frank symptoms of the disorder themselves emerge from a complex (and plausibly in-deterministic) genetic-development-environmental interplay (6–9).

Lack of understanding of causative pathways (in addition to other factors such as phenotypic heterogeneity) is a limiting constraint on efforts at prevention, early intervention, and/or treatment of schizophrenia (10). Yet, as is evident in other fields of medicine (most notably cardiology and oncology), successful medical treatment and management does not necessarily depend on identifying deterministic causal pathways toward disease. Rather, understanding of vulnerability measures is sufficient, as long as there is a clear characterization of the pathophysiological *mechanisms* underlying the disease. A grasp of basic biological mechanisms drives the development of targeted therapies while simultaneously providing objective biomarkers that can inform treatment efficacy (11, 12). Psychiatry has long been criticized for focusing on emergent *effects* of disease, while having an insufficient focus on understanding *mechanisms* of disease. As a result, many emergent effects are insufficiently constrained by analytic approaches designed to explicitly characterize mechanisms. The argument is that these limitations essentially limit nosology and treatment (13). Here, we reiterate the widely held view that the mechanisms of diseases like schizophrenia will most profitably be understood by focusing on, to put it simply, how the brain is “not working.” More pertinently, we accept the position that the current state of acquiring and modeling brain signals suggests that brain–behavior relationships may best be understood from the position of *macroscopic* brain network interactions (14), a scale that may most proximately map cognitive and sensorimotor function to its underlying correlates. We are aligned with the position that these investigations are a matter of discovery, and subscribe to the distinction between “true” models of brain function (whatever those may be) and “likely” models of brain function (15), where the former are theoretical constructs, whereas the latter are empirically discoverable from neuroimaging data. We will assert that these ontological subtleties are of direct (and not merely academic) relevance to the study of schizophrenia: if understanding brain mechanisms subserving normal function is a process of discovery, by corollary, understanding these in the context of schizophrenia is also a process of discovery. Moreover, this process is constrained by quantitative models, and the assumptions therein, that are applied to any class of neuroimaging data. Thus, understanding dysconnection in schizophrenia is an *inference*, and the class of models applied will provide constraints on the type of inference that can be drawn (16).

What follows is a compendium of previously advanced ideas motivating schizophrenia as a syndrome of functional disconnection, or more accurately, “dysconnection”: this is generally understood as the abnormal integration of signals across the brain (17). These ideas are not novel to this review and, in fact, have remained in place since the earliest conceptions of schizophrenia itself. However, over time the idea of dysconnection, once a general construct, has been systematically hewn to a point where it can be seen as having clear bases in translational neuroscience (18–20). We willfully restrict our scope to the approaches used in the analyses of functional magnetic resonance imaging (fMRI) signals. Thus, while much of what we review strictly extrapolates

to fMRI studies in schizophrenia, many of the quantitative approaches that we review are not specific to fMRI time series analyses.

Much has been written about fMRI, its neurophysiological bases, the relative advantages of the signal, and its limitations for assessing brain function (21–24). Moreover, the comparative merits and demerits of different imaging techniques for discovering connection and disconnection in the brain is a viable topic but beyond our scope (25, 26). As a technique, fMRI is neither mercurial nor worthless; rather the technique provides access to credible signals that under appropriate analytic constraints can tell us something interesting about how the brain works.

ORGAN SYSTEMS AND PHENOTYPES

As understood from the early origins of psychiatric taxonomy, the body’s only organ system of direct relevance to schizophrenia (and indeed all psychiatric illnesses) is the brain, and the brain’s *functional* properties were seen as the most salient in this regard (27, 28). Psychiatric phenotypes have always fundamentally been defined by behavioral abnormalities (29). If the “mind–body” problem is narrowly defined as the problem of understanding how mental states emerge from the physical states of the body (30, 31), then understanding pathophysiological mechanisms underlying schizophrenia is a special case of the mind–body problem. That is, if “neural” processes drive normal behavior, then abnormal behavior (i.e., such as those observed in schizophrenia) must result from abnormal neural processes. A challenge then is to identify abnormal neural mechanisms that lie in a straightforward relationship with the phenotypic characteristics of schizophrenia itself. Given that the dialectic regarding the mind–body problem remains active, it is self-evident that understanding neural mechanisms associated with schizophrenia is a non-trivial problem.

The term “dysconnection” itself can be construed as having at least two somewhat distinct meanings. In one, the term represents the sense in which Bleuler and Kraepelin thought of schizophrenia: a kind of “splitting of the mind” (32, 33). In all likelihood, this meaning carried only a vague relationship to neurobiological considerations. In his writing, Bleuler clearly related the idea to loss of cohesion in intellectual faculties, but the work of these great neurologists predated modern neuroscience, and thus they were not privy to the multitude of experimental methods and theoretical ideas available now. Indeed, in their time, understanding of brain function was largely grounded in phrenological/localizationist theories wherein the outputs of individual brain regions, emergent as their “functions,” were assumed to map onto anatomical structures in relatively direct ways (34, 35).

A second, more literal sense of dysconnection (or “disconnection” as originally proposed), is quite explicitly neurobiological: in this view, schizophrenia is an emergent behavioral phenotype resulting from profound alternations in the connectivity of the brain’s anatomic and functional pathways (18, 36, 37). Just as structural connectivity loosely constrains the brain’s functional architecture (38, 39), impaired anatomical connectivity most likely offers loose constraints on the inability of the brain in schizophrenia to integrate functional signals across regions in

both non-task and task-active states (40, 41). It has been asserted that normal perception and cognition rely on a cortico-cortical phase synchronization mechanism that operates in conjunction with reentry to provide context for local cortical computations by way of inter-areal interactions (42). This mechanism normally provides a balance of integration and segregation to complex dynamic cortical networks, and schizophrenia is likely to be marked by a shift toward segregation such that local cortical areas express their information content without benefit of the context normally provided by interaction with other areas (43, 44). A modern view of the dysconnection syndrome constitutes a program for neurobiological discovery and offers the promise for arriving at fundamental insights on brain mechanisms that mediate the emergence of the illness.

FROM LOCALIZATION TO NETWORK FUNCTION: NEUROLOGY AND NEUROPSYCHIATRY

The localizational model was in part driven by the technologies available for collecting data from the brain. In essence, the model relied on systematic analyses of functional loss in patients with localizable neurological lesions (45). Lesion-based models were characterized by rudimentary system's-based approaches toward brain-behavior relationships, the Wernicke-Geschwind model of speech comprehension and production (46) serving as a good example. However, their accuracy was compromised by impoverished data and, more fundamentally, by the untenable (yet implicit) assumption that brain regions existed in fairly specific one-to-one relationships with overt behavior (34). Among others, Wernicke and Luria should be given credit for creating a hybrid model that distinguishes between elementary functions expressed by individual brain regions and complex functions that are properties of distributed systems of brain areas (42). Thus, as we understand it now, early neurology turned out to be an impoverished framework for understanding both the complexities of normal brain function and the complexities of disorders like schizophrenia.

The “neuron theory,” largely motivated by the work of Santiago Ramón y Cajal (47), emerged in the late nineteenth century. The theory experimentally developed the idea of neurons as functional units, the extracellular outputs of which reflect basic properties of behaviors (both simple and relatively complex). The neuron theory has been a (perhaps *the*) *sine qua non* of modern neuroscience. The explosion of single unit extracellular recordings has provided a wealth of insight into the complex response properties of single units and the extent of (particularly sensorimotor) function that these responses explain (48, 49). Nonetheless, theories of structure-function relationships based on single unit recordings are ultimately subject to the same ontological limitations as phrenology (50). Cognition and behavior (normal or pathological) are far too complex to reduce to single brain regions, let alone units. Indeed, as connectivity is a basic property of neurons (which connect through axons and synapses), connectivity studies are a natural direction for neuroscience. If anything, neurological models, neuron theories, and lesion studies provide evidence of

only one of the multiple organizing principles of brain function at the macroscopic scale: the principle of relative specialization of function. This principle suggests that brain regions are more likely to sub serve one class of functions than another (51, 52). Relative specialization is a question of degree. There is little about the brain that is strictly “categorical.” Rather, it is more likely that the degree of specialization is relatively strong in sensorimotor or modality specific regions (“unimodal” regions), but relatively weak in regions involved in sensorimotor integration and “higher” behaviors [or “heteromodal” regions (53, 54)]. Clearly then, a different and parallel principle of functional brain organization is needed that more directly speaks to the emergence of complexity, and by corollary, the emergence of complex disorders of the brain.

If we treat the single neuron theory as the first revolution in modern neuroscience, the application of complex systems theories, and the operationalization of these analytic frameworks for understanding the brain must constitute the second (55). Ludwig von Bertalanffy's work on general systems theory continues to reverberate in modern neuroscience (56) as the drive to explain processes in terms of interactions between the components of a system gains force (57). This is particularly pressing in terms of understanding how cognitive ontologies arise from the brain (39, 58). The functional integration of signals across brain regions presents itself as a parallel organizing principle of brain function (52, 59), and if complex cognitive ontologies arise from brain network interactions, and schizophrenia itself is in part defined as a “cognitive” illness (10), then it is reasonable to assert that the disorder results from impaired functional integration of signals across brain regions, i.e., brain network dysfunction.

The history of neuroscience is characterized by a synergistic relationship between methods and theory (60, 61). Theoretical advances in the neurobiology of schizophrenia have been crucially driven by developments in *in vivo* whole brain functional neuroimaging, in particular fMRI. fMRI is based on complex hemodynamic spatiotemporal signals (62, 63) that themselves lie at the apex of a series of complex neuronal (and presumably neurochemical) processes; these processes exist in uncertain relationships with the overt fMRI signal (22, 61). Thus uncovering brain network function and dysfunction from fMRI time series data has been termed a process of “reverse engineering” and “network discovery”; what must be engineered or discovered are the hidden states of brain function that give rise to fMRI signals (55, 64). If understanding brain network interactions is a process of quantitative discovery, then understanding of what the inferred processes are must be grounded in quantitative models applied to fMRI time series data (65).

MODELS FOR INFERRING NETWORK FUNCTION: WHAT CAN BE INFERRED?

Distinct classes of analytical techniques have been fruitfully applied to fMRI time series data (38, 66, 67), and much of what follows is a synthesis of published and influential reviews. Our goal is to selectively sample and distil for the reader quantitative bases of analytic methods, and reveal what they render as knowable about brain connectivity and, by extension, dysconnectivity.

Ultimately, connectivity is an inference that is based in the quantitative models used to assess it. This inference inherits the advantages and limitations of the models applied to discover these patterns in overt fMRI signals. In understanding brain network dysfunction in schizophrenia (or any psychiatric condition) exposure to the fundamentals of the analytic methods is a necessity for understanding what is being modeled (and ultimately inferred).

Functional connectivity (FC) and effective connectivity (EC) are two broad classes of analytic methods for assessing connectivity. FC generally refers to the statistical relationship (in the time domain) of two spatially distinct signals. In fMRI data, these analyses typically constitute calculations of bivariate temporal correlations usually at zero-lag thus ignoring potentially useful information about timing relations between BOLD time series drawn from distinct regions of interest, wherein strongly correlated or anti-correlated regions are “functionally connected” (68). In general, co-variations in time series signals are heavily used in assessing functional relations between elements of complex systems (69) and provide a useful though limited framework for network discovery. In the context of fMRI data, these methods have been noted for some limitations. First, they are sensitive to the variability of hemodynamic response functions (HRFs) across brain regions. Second, due to the limited temporal information in the signal, they generally must rely on “zero-lag” analyses, thus ignoring potentially useful information about timing relationships between BOLD time series. Moreover, they lack functional transitivity, in that the techniques are insensitive to divining functional relationships between regions that are correlated with a mediating time series (55, 70, 71). Coherence is a complementary measure of FC that estimates linear time-invariant relationships between multiple time series even at phase delays. As a spectral analog of bivariate correlational analyses, coherence accounts for phase relations in the cross-correlational function (while limited by the bandwidth of the hemodynamic response) (72). Finally, we also note the value of Wiener–Granger Causality (73), a method for inferring functional relationships between regions based on temporal predictability between time series. Granger Causality has also been referred to as directed FC (dFC), distinguishing it from undirected FC (uFC) methods that depend on correlative statistics (74, 75).

The standard notion of EC is that it captures the effect that one neuronal population exerts over another (regardless of the scale at which these interactions are assessed) (52). More specifically, true EC depends on capturing the relatively precise timing relationships between *neuronal* populations (76), thus depending on the applied models to capture the temporal dynamics of neuronal populations. Implicit within the definition of EC is that these techniques explicitly seek to model “causative” relations between brain regions and that they depend on generative network architectures of the brain. A causal relationship is both an ontologically different claim than simple statistical covariation, but importantly therefore also includes information about the direction of influences between regions (77). Moreover, EC entails a notion of estimated “coupling”: that is, the determination that a causal influence exists between neuronal populations fundamentally relies on constructing a *generative model* of

that coupling. This means that understanding EC is a question of model comparison: what model of the brain is most likely to have generated the observed data, where each model itself constitutes a hypothesis of brain function in that context (78). The implicit and explicit assumptions of each approach and the models implemented therein exert constraints in inferences regarding brain network connectivity, and by extension, about dysconnectivity in schizophrenia. Beyond this, modeling can help unravel the physiological mechanisms underlying causal influences in the brain. That is, what is the effect exerted by the neurons in one brain area on those in another area. It is plausible that inter-regional interactions are modulatory, shaping the activity generated by the internal dynamics of the local area (42). To evaluate this hypothesis, we need to be informed by (a) the location on the recipient neurons of the synapses coming from transmitting neurons in other areas and (b) the neurophysiological effect of inter-regional influence (e.g., how does it change the sub threshold membrane potential?). Moreover, model *structure* plays an important inferential role. The likely generative model for schizophrenia may be different than that for controls, and the differences in model structure may provide valuable information regarding the architecture of the disease (41). These issues are revisited later.

UNDIRECTED FUNCTIONAL CONNECTIVITY TECHNIQUES BASED ON BIVARIATE CORRELATIONAL APPROACHES

Bivariate correlational approaches toward fMRI time series data are by far the most commonly applied measure of assessing temporal relationships between regions of interest. These methods make weak assumptions regarding functional transactions between regions, are frequently used in an exploratory manner, and do not provide measures of coupling in the manner of other techniques (see more on this aspect below) (75). Generally considered, this class of FC analysis mines statistical dependencies in the time domain (for fMRI data) between disparate time series. This approach is represented in:

$$\rho_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y}$$

In this generic form, the correlation coefficient, $\rho_{x,y}$, of two independent time series x and y , is equal to the covariance of x and y normalized by the product of the SDs, σ , of both signals. FC is an *emergent* statistical property of inter-relationships between time series, such as provided by the fMRI BOLD signal. Crucially, FC analyses have not usually relied on biophysical models linking neuronal with hemodynamic responses (79), which means the notion of “coupling” in uFC analyses does not extend beyond the statistical realm. Thus, as with any emergent statistical result, the only “model” of function tested by uFC methods is the null model, that is, testing against an absence of significant correlation between brain regions. As a result, these methods make few assumptions about temporal context, the temporal scale, or resolution of any putative underlying process.

The weak assumptions relating to processes, however, also confer some advantages for this class of FC measures. There is greater statistical reliability with shorter time series than measures that make stronger assumptions or that require more precisely modeled time series (as will be seen with Granger Causality techniques). Moreover, weak assumptions also mean that uFC analyses are relatively robust to contributions of filtering or slice-time correction applied to fMRI data that can artificially disrupt the fine-scale temporal structure of the signal, thus producing spurious causality (we discuss the relationship between temporal information and causal inference in detail below).

uFC TECHNIQUES BASED ON COHERENCE

Coherence is a complementary approach to time-domain uFC that has enjoyed widespread use with EEG data, but has been generally under-utilized for fMRI analysis, largely because limited temporal information in the fMRI signal preempts complex spectral analyses of fMRI signals. Coherence is a frequency-domain measure of how well one signal linearly predicts a second in a time-invariant fashion. The most common approach defines coherence as the magnitude of the cross-spectrum of two signals, x and y , normalized to the power spectra of each signal. Specifically, magnitude squared coherence, Coh_{xy} , at a given frequency, f , can be defined as

$$Coh_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)}$$

where P_{xy} is the cross-spectrum, and P_{xx} and P_{yy} are the power spectra of both signals. As with correlation, the metric is scaled ($0 \leq Coh_{xy}(f) \leq 1$), such that 0 represents no linear relationship between the two signals, and 1 represents the ability to perfectly predict one signal from the other. In order to reduce the variance and edge artifacts that can be introduced by windowing data in the time domain prior to a Fourier transform, coherence is often calculated using Welch's modified periodogram averaging method (80). An estimation of the interregional coherence can then be calculated for each resulting frequency bin or averaged over the frequency range inhabited by the hemodynamic response function (HRF), typically defined as 0–0.15 Hz (70).

The time-invariant property of coherence allows the measure to assess relationships between time series beyond the zero-lag constraint. As previously noted, variations in the shape of the HRF along with time delays may offset the temporal progression of two related BOLD responses in functionally related brain regions. This property of fMRI signals presents a challenge for the sensitivity of correlation analyses, because correlations will decrease toward 0 as a function of increasing lag between two otherwise synchronized signals. For example, if two brain regions are co-modulated by a task, but with a time delay or with different hemodynamic responses, they may be synchronized with a phase lag. This can be generalized as the relationship between a sine function, $\sin(t)$, and a second identical, but phase-shifted sine function, $\sin(t + \theta)$. As the phase shift, θ , progresses from 0 to $\frac{\pi}{2}$,

the zero-lag correlation decreases to 0, then further to -1 as θ approaches π . However, a linear relationship still exists between the two signals, and coherence between the two sine waves will remain at 1. In other words, coherence allows for phase shifts or temporal lags when scoring the FC. Thus, while both zero-lag correlation and coherence are able to capture the relationship between the signals when the phase lag is near 0 or π , lags around $\frac{\pi}{2}$ are lost to zero-lag correlation [for more details on this method, see Ref. (71, 80)].

Spurious non-zero values of coherence can arise in the analysis of physiological time-series data simply due to the spectral properties of the signals (81). An effective method for correcting this bias is the use of surrogate data sets (81–83). Here, the time series data are shuffled, thus disrupting the phase relationships but preserving the statistical distribution of the spectral content. After multiple iterations, the mean and SD of the surrogate results can be compared to the experimental results. Any values of coherence exceeding the 95th percentile of the surrogate data distribution can be considered significant.

Correlative analyses are undirected by definition – hence the label “undirected FC” above. An interpretational limitation is that these methods are agnostic regarding directional influences between network nodes. Undirected connectivity/coherence analyses have a different footprint in the analyses of electrophysiological signals where within- and inter-regional coherence can be assessed at multiple frequencies (84, 85), each of which reflect somewhat distinct functional properties of cortical function. The temporal resolution of fMRI data does not afford this luxury. Nonetheless, there has been an exuberant profusion of undirected FC techniques applied to the resting state fMRI. Combined with graph-theoretic methods, these applications have provided insight on network disorganization in schizophrenia (40, 44). A valuable extension of this work would be in understanding how these altered network hierarchies in schizophrenia are expressed in disconnection in a task-active state. However, such extension would require integrating multiple areas of focus, including within-subjects acquisition of resting *and* task-based data, and the implementation of multiple techniques for estimating connectivity (outlined herein). Nevertheless, we suspect that the value of uFC analyses alone in inferring disconnection in schizophrenia is limited. The applied statistical model for uFC is relatively impoverished and identifies emergent statistical properties of fMRI signals [see Ref. (16) for a more comprehensive treatment of these questions]. These emergent statistical properties are removed from biophysical models linking accumulative neuronal with hemodynamic responses (86) and therefore may be distant from mechanisms of brain function.

Understanding directional relations between network constituents is important for multiple reasons. While various brain areas have reciprocal structural connections, it is highly unlikely that directional relations will simply reflect structural connections. Moreover, for a variety of reasons, structural connectivity only offers loose constraints on the functional integration of signals. In terms of functional organization, it is very likely that the direction of information flow in the brain is of critical importance for organizing cognitive functions and consciousness. This has been

demonstrated from numerous studies of connectivity in both EEG and fMRI (87, 88). Moreover, directional effects are also important in the context of brain network hierarchies (53, 54). Control regions of the brain, including the dorsal prefrontal cortex and the anterior cingulate cortex enjoy higher hierarchical status within the overall system (74, 89, 90), suggesting that their functional transactions are likely to be asymmetric (91). Capturing these asymmetries will prove highly valuable, particularly in schizophrenia, which has frequently been characterized as a disorder of cognitive control (92–94). These considerations motivate quantitative FC methods that explicitly attempt to capture directional interactions between network constituents in health and schizophrenia. We next consider two directed FC methods, Granger Causality (73) and psychophysiological interaction (PPI) (95, 96).

FC TECHNIQUES BASED ON DIRECTIONAL APPROACHES

Psychophysiological Interaction

Since its introduction in 1997, PPI has constituted a widely used approach to directed FC (95). PPIs are constructed by extracting a time series from a seed region of interest and multiplying its activity with a stimulus function or regressor encoding the psychological context (95). This computation generates a regressor term that is used to capture variance in the time series of target voxels, as explained by the seed region, within the context of the task. Technically, signals that are highly predictable will produce a significant PPI effect but PPIs are readily distinguished from correlative methods. This follows because they test for second order dependencies. In other words, they test for a linear dependency of activity in the target region on activity in the seed or source region that itself depends upon another (psychological) variable. It is this high order, or interaction, effect that breaks the symmetry and endows PPI analyses with a directed nature. Strictly speaking, one could argue that PPIs reflect a simple (GLM) model of EC. However, we associate a PPI analysis with the inference that there are statistically significant second order dependencies; namely the interaction. As such, we will treat PPI analyses as a form of directed FC (i.e., statistical dependence).

The GLM approach to assessing PPIs provides a potentially more nuanced framework for modeling the time series data as it allows the model to co-vary out confounds. This is accomplished *via* a point-wise multiplication of the seed time series with a stimulus function and the inclusion of various sources of noise in the model (96). This product time series is then the interaction between the BOLD time series and the psychological task – the eponymous PPI. The interaction time series, along with both the task model and HRF time series, can then be used as regressors in the GLM, which separates the variance in the target signal associated with the psychological task, the HRF, and the interaction between the two signals (the psychological regressor and the response in the seed region). The equation below captures the directional bases of PPIs:

$$y_i = ay_0 + b(y_0 \times u) + cu + X\beta$$

The above is readily distinguishable from the typical GLM applied in activation models by the presence of the asymmetric interaction term $(y_0 \times u)$, in which regressing $(y_0 \times u)$ on y_i is asymmetric with regressing $(y_i \times u)$ on y_0 (55).

As with all modeling of fMRI signals, PPIs constitute an *a priori* conceptual model of brain function. The implicit model is that *contextual interactions* between seeds and targets can be characterized within a statistical framework. In this context, the choice of seed and the psychological context are free parameters of the model, and these choices must be well motivated by prior knowledge regarding task characteristics and the putative network profiles of the seed in an integrative network (94, 95, 97, 98). The simplicity of the PPI framework is advantageous as it affords rapid exploration of network profiles in normal and clinical populations (and differences between them). Nevertheless, this simplicity is also a limitation because comprehensive network interactions are rarely subsumed by pairwise interregional interactions. Moreover, PPIs are not defined by biophysical models linking neuronal with hemodynamic responses, and therefore they do not provide measures of neuronal coupling, limiting their neurobiological interpretation (95, 96).

Granger Causality

In treating the brain as a complex system, computational neuroscience has successfully coopted analytic tools that were initiated for other disciplines, where many data properties are shared with fMRI data – particularly the idea that functional aspects of the system interactions are “hidden” in time series data. Notable examples from physics and electrical engineering include time-frequency analyses, graph theoretical approaches, and information theoretic measures. GC is a measure of directed FC, which has its roots in the analysis of economic data (99). Since its import into the field of neuroscience, GC has been used extensively in estimating directed connectivity relationships in multiple modalities of brain imaging including EEG, MEG, and fMRI (73). GC can, in brief, be described by:

$$X_t = \sum_{j=1}^m \alpha_{2j} X_{t-j} + \sum_{j=1}^m \beta_{2j} Y_{t-j} + \epsilon_{2t}$$

$$X_t = \sum_{j=1}^m \alpha_{1j} X_{t-j} + \epsilon_{1t}$$

$$GC_{Y \rightarrow X} = \ln \left(\frac{\Sigma_1}{\Sigma_2} \right)$$

where, $\Sigma_1 = \text{Var}(\epsilon_{1t})$ and $\Sigma_2 = \text{Var}(\epsilon_{2t})$. In GC, the estimated dFC depends on the model order and the estimated time lag between modeled time points in the two signals. Accordingly, appropriate specification of the time lag between different observations and the order or number of past observations included in the auto-regression model above are crucial for properly testing for GC influences. Generally, the appropriate time lags and model orders are unknown (although they can be estimated using procedures based upon mutual information and Bayesian model comparison). This means the ability of GC to make inferences

about FC – based on relatively slow dynamics – is potentially challenging [see Ref. (100, 101)]. It should be noted that correlation also suffers from a dependence on time lag, and important correlations may be missed by over-reliance on the lag-zero value.

In general, GC evaluations may cover a range of time lags up to a maximum value in order to search the parameter space for the correct lag. This approach has been demonstrated for transfer entropy, an information theoretic extension of GC. Lee et al. (87) define the information flow between two nodes as the maximum normalized transfer entropy found by scanning the time delay parameter space. This method has been successfully applied to electroencephalograph data in humans (87), as well as electrocorticograph data in rodents (102), allowing researchers to track changes in cortical information flow across changes in states of consciousness. However, given the limited temporal resolution of fMRI, the model order rarely extends beyond one time step.

Because GC relies on the temporal progression of time series to estimate causality, it is crucial that the ordering of the time series data remain unmodified. The convolution process used in finite- and infinite-response filters alters each data point based on both the previous data points and those succeeding it. After filtering, then, the value at any given time has been influenced not only by its past but also by its future, thus undermining the basic assumption of causal inferences. Convolution with a HRF can also affect the structure of time series, notable because as referred to earlier, HRFs can differ across brain regions. Several lines of work somewhat mitigate against these concerns. There is evidence that GC is invariant to HRF convolution (103), and as noted earlier, assessing directional asymmetries (i.e., direction is used as a condition of interest in the GLM) of GC coefficients between regional pairs can better constrain the interpretation of causal effects (104).

Notwithstanding the directed nature of the applied models, as with PPI, GC is limited by a lack of adequate bases in physiological underpinnings. The dynamics of interacting neural populations is considerably removed from the signals recorded by fMRI and subsequently entered into a GC calculation, and there is no extant physiological model to bridge it. Nevertheless with appropriately constrained research questions and experimental designs, GC (as with PPIs) provides a quantitative characterization of directional pairwise interactions between constituents of brain networks. These insights can provide meaningful perspectives on the nature of dysconnection in schizophrenia.

EFFECTIVE CONNECTIVITY TECHNIQUES

The term “effective” connectivity is the source of much confusion, but can be more clearly understood from the perspective of interactions between constituents of a complex system. From a mechanistic perspective, EC has been defined as the influence that one neural system exerts over another (76). EC models thus attempt to embody dynamic and timing relationships between system constituents, typically using elements of control theory (52). Because there is no reasonable sense of a “true” and veridical model of brain function that can be wholly derived from observed data (15), EC relies on a method for evaluating *competing* model architectures for a set of network nodes, any

of which is a plausible generator of the observed data, but have inherently differing likelihoods of having done so (78). This competitive (and ultimately Bayesian) framework is essential for the process of model discovery and is largely absent from previously considered techniques of directed and undirected FC (though some elements are present in the evaluation of directional asymmetries using GC). These motivations for divining neuronally plausible network-based mechanisms are by themselves insufficient unless a mapping from neuronal responses to the generated hemodynamic response is implemented (86). Dynamic Causal Modeling (DCM) incorporates a Bayesian framework for network discovery while using biophysical models relating neuronal to hemodynamic responses.

Dynamic Causal Modeling

Dynamic causal modeling was introduced as a seminal framework for discovering mechanistic brain network function from fMRI (and other) data (105) and has subsequently received significant methodological inspection of its biophysical and probabilistic bases, and its reliability (106, 107). In its original conception, DCM represented the brain as a bilinear system (a lower order approximation for non-linearity) in which the inputs are the experimental conditions and the outputs are the hemodynamic response measured using fMRI. Because DCM is explicitly interested in modeling dynamics and changes in these dynamics in response to inputs, elements of control theory are incorporated where changes in network states are modeled using the following state differential equation:

$$\frac{dx}{dt} = \left(A + \sum_{j=1}^m u_j B^{(j)} \right) x + Cu$$

To be clear, this equation describes the dynamics of “network states” rather than of any physiological metrics. In other words, this is not an explicit physiological model. The state equation represents three differentiable components with large distinct bio-psychological extensions: here, A represents the matrix of endogenous coupling between brain regions. Simply speaking, this can be construed of as the hypothesized functional connectome underlying the evaluated model. The connections within any model represent a hypothesis on the pattern of connectivity. How the model is “wired” depends on a combination of priors that include reliably known connective pathways and/or explicit hypotheses for discovering which pathways may or may not exist independent of task-induced experimental changes. The variable $B^{(j)}$ represents the modulatory response in the network connections due to changes in experimental conditions u_j . Finally, C represents the direct driving input on particular regions as induced by the experimental conditions.

Network discovery with DCM relies on the identification of generative network architectures with the highest evidence given the observed fMRI data, thus testing hypotheses on an *a priori* defined model space. Thus, the space comprises neurobiologically plausible competing models, each representing hypotheses on the connective-architecture of the investigated neural system (78, 108, 109). Therefore, rather than using traditional goodness-of-fit metrics to assess the viability of an individual model, DCM

relies on evaluating these multiple neurobiologically plausible competing network models, wherein across competing models, specific network connections may be permuted, and these permuted connections specifically serve as hypotheses. Thus, across models, connections can be constrained or informed by known properties of neuroanatomy (110) or may be permuted with some agnosticism regarding the specifics of the underlying anatomy itself. More importantly, the method provides a plausible approach toward understanding how brain networks “work,” by incorporating principles of both relative specialization and functional integration. By corollary EC methods provide a compelling context for understanding cases (such as schizophrenia) in which the brain does “not work” (111).

CHALLENGES FOR INFERRING CONNECTION AND DYSCONNECTION

Understanding *causal* antecedents of disease have historically driven the promise of successful medical intervention or pre-emption. This concept of *deterministic* causality is a fundamental assumption in medicinal discovery and practice. Illnesses have causes that can be determined and addressed. These causes may cut across biological, environmental, and epidemiological levels (112), yet it is commonly assumed that the causative pathway is, in fact, deterministic. Moreover, many medical successes have benefit from the relative simplicity of the structure–function relationships within the organs of relevance. For example, advances in cardiac or pulmonary treatment benefit from the relatively straightforward relationships between the structures of the heart or the lung, and their expression in function. Unfortunately, the brain proffers no simplicity in this regard. Rather, its structure (to the extent that it is fully knowable) provides relatively light constraints on its emergent functional interactions, and even more indeterminacy with respect to how overt behavior arises. This “degeneracy” is a significant challenge to assessing the brain’s structure–function relationships (35) and degenerate structure–function mappings (i.e., many to one mappings) can also be observed in the relationship between functional and effective connectivity. In other words, there are many connectivity architectures of EC that can produce the same FC. We will consider an example based upon a common source below. Technically, this means that inferring EC from FC can be an ill-posed problem that necessarily calls for prior constraints and (abductive) inference, as we have carefully intimated in the Introduction.

For instance, should connectivity between two regions be established *via* a statistical model, the conclusion that the two regions have a direct relationship can always be undermined by the hypothetical existence of a third, undetected member of the system, though this possibility can be somewhat mitigated by knowledge of the anatomical connectivity structure of the areas under consideration (thus, region C may only be a candidate driver if it is anatomically known to project to regions A and B). Thus, a bivariate correlation between regions A and B, may be driven by (1) a functional or causal relationship in which A causes B and/or B causes A, or (2) regions A and B are both modulated

by a third region, C, such that A and B are synchronized without directly interacting. If there is a different time delay between C and A than between C and B, this can even appear as a phase-lag relationship between regions A and B thus presenting as a directed FC relationship. However, given that methods such as GC are based on predictability, they can be used to establish a driving role for area C (113). Thus, demonstrating a directed relation based on any measure of predictability is superior to phase-lag measures in terms of what can be inferred. Nevertheless, although multivariate GC can account for a third time series and its influence on the system, this can only be modeled when the third time series has been identified and is measurable. If the third node in the system is unknown, then contribution toward the inference of connection remains unknown. These limitations are a particular example of the general problem of causal inferences in brain networks, and as has been forcefully argued and extensively discussed (114, 115), deriving *deterministic causal inferences* regarding brain network interactions may be a fundamentally untenable exercise. Not only might the etiology of schizophrenia remain obscure, but even the inference of brain network mechanisms that might inform the proximate causes of the illness may suffer from fundamental challenges. The potential contribution of hidden or latent nodes is, in principle, not a problem for models of EC like DCM. This is because one can use Bayesian model comparison to evaluate the probability of a hidden common source – by comparing models with and without hidden nodes.

WHAT ASPECT OF DYSCONNECTION IN SCHIZOPHRENIA CAN BE UNDERSTOOD FROM THE ANALYSES OF fMRI SIGNALS?

Figure 1 summarizes the methods visited in this overview, arranged in a two-dimensional conceptual space. The space carves out quadrants within which the methods are defined by their directionality (directed vs. undirected network interactions) and the relative “strength” of the methods (functional or effective connectivity), where “strength” can reasonably be construed of as the degree to which the parameters of connectivity models can be linked to biophysical processes or computation in real neuronal networks.

This is a parsimonious representation, yet provides a panoramic (and somewhat self-evident) overview of the different aspects of network dysfunction that can be inferred from each class of analytic approaches applied to fMRI time series data. Though we emphasize that the axes do not approach metric properties, we use weak ordinality to arrange the methods within the space to allow some contemplation of their strengths and weaknesses. For example, directed analyses may provide more interpretational value than undirected analyses, and EC approaches may, in principle, be more desirable than FC approaches. A crucial dimension not represented is tractability of implementation – both in terms of designing analyses and computationally implementing them. We note that this tractability is unevenly distributed within this space, yet is an issue of concern in the search for inferring dysfunction in schizophrenia. We also note

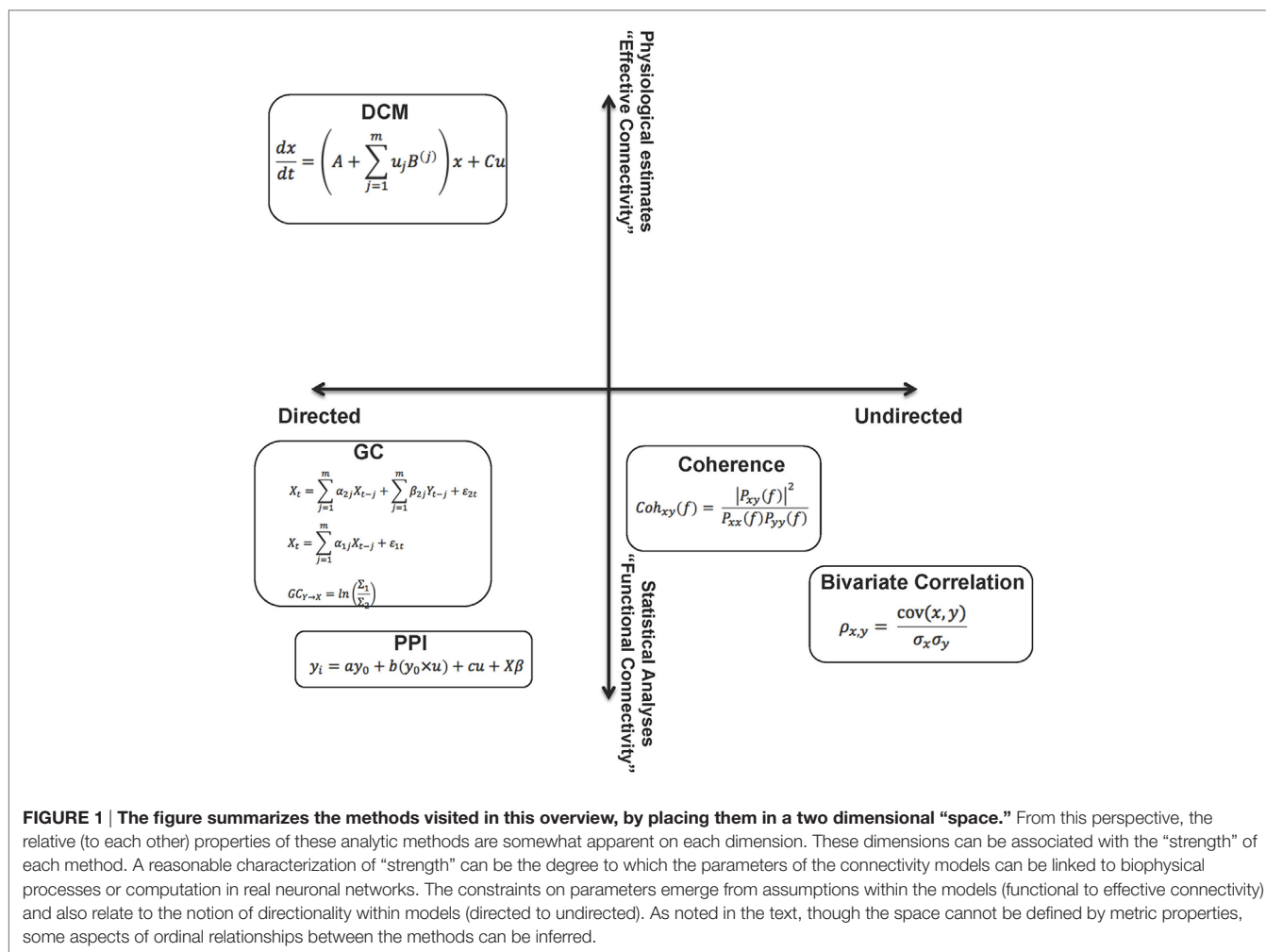


FIGURE 1 | The figure summarizes the methods visited in this overview, by placing them in a two dimensional “space.” From this perspective, the relative (to each other) properties of these analytic methods are somewhat apparent on each dimension. These dimensions can be associated with the “strength” of each method. A reasonable characterization of “strength” can be the degree to which the parameters of the connectivity models can be linked to biophysical processes or computation in real neuronal networks. The constraints on parameters emerge from assumptions within the models (functional to effective connectivity) and also relate to the notion of directionality within models (directed to undirected). As noted in the text, though the space cannot be defined by metric properties, some aspects of ordinal relationships between the methods can be inferred.

that the motivation for this review relates to schizophrenia, yet the extensions are general across psychiatric illness (though the networks of focus within the brain may be idiosyncratic to the phenotypes of interest).

Is schizophrenia itself *tractable*? It is unclear whether all aspects of the etiology of this complex condition are knowable, yet inferring dysconnection in schizophrenia is a special case of inferring brain network function. In that sense, inferring the dysconnection syndrome is perhaps no more or no less tractable than understanding how macroscopic brain network interactions can be related to other overt or covert behaviors. We suggest that if fMRI has told us anything of significant value, it is that macroscopic brain network dynamics expressed at the scale of seconds can be successfully modeled to infer aspects of brain function. Theory and technique now offer avenues for inference and discovery that did not exist even in the recent past. The methods covered herein (and others) offer the prospect of inference and discovery that suggests the promise of significant mechanistic understanding of schizophrenia.

AUTHOR CONTRIBUTIONS

VD directed the writing of the manuscript and contributed the principle conceptual direction. SB provided important conceptual insight and technical language particularly relating to Granger Causality models. BS provided insights relating particularly to functional connectivity methods.

FUNDING

Preparation of this work was funded by support to VAD: a Career Development Chair from Wayne State University, the Thomas Gershenson Distinguished Faculty Fellowship from Wayne State University, the Lyckaki-Young Fund from the State of Michigan, the Prechter Family Bipolar Foundation, the Children’s Hospital of Michigan Foundation, the Cohen Neuroscience Endowment, and the National Institute of Mental Health (MH 59299).

REFERENCES

- Petronis A, Paterson AD, Kennedy JL. Schizophrenia: an epigenetic puzzle? *Schizophr Bull* (1999) 25:639–55. doi:10.1093/oxfordjournals.schbul.a033408
- Schultz SK, Andreasen NC. Schizophrenia. *Lancet* (1999) 353:1425–30. doi:10.1016/S0140-6736(98)07549-7
- Keshavan MS, Tandon R, Boutros NN, Nasrallah HA. Schizophrenia, “just the facts”: what we know in 2008 Part 3: neurobiology. *Schizophr Res* (2008) 106:89–107. doi:10.1016/j.schres.2008.07.020
- Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL. The human disease network. *Proc Natl Acad Sci U S A* (2007) 104:8685–90. doi:10.1073/pnas.0701361104
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* (2014) 511:421–7. doi:10.1038/nature13595
- Lewis DA, Levitt P. Schizophrenia as a disorder of neurodevelopment. *Annu Rev Neurosci* (2002) 25:409–32. doi:10.1146/annurev.neuro.25.112701.142754
- Keshavan M, Kennedy JL, Murray RM, editors. *Neurodevelopment and Schizophrenia*. Cambridge: Cambridge University Press (2004).
- Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, “just the facts” what we know in 2008. 2. Epidemiology and etiology. *Schizophr Res* (2008) 102:1–18. doi:10.1016/j.schres.2008.04.011
- Howes OD, Murray RM. Schizophrenia: an integrated sociodevelopmental-cognitive model. *Lancet* (2014) 383:1677–87. doi:10.1016/S0140-6736(13)62036-X
- Tandon R, Nasrallah HA, Keshavan MS. Schizophrenia, “just the facts” 4. Clinical features and conceptualization. *Schizophr Res* (2009) 110:1–23. doi:10.1016/j.schres.2009.03.005
- Javitt DC, Spencer KM, Thaker GK, Winterer G, Hajos M. Neurophysiological biomarkers for drug development in schizophrenia. *Nat Rev Drug Discov* (2008) 7:68–83. doi:10.1038/nrd2463
- Goff DC, Romero K, Paul J, Mercedes Perez-Rodriguez M, Crandall D, Potkin SG. Biomarkers for drug development in early psychosis: current issues and promising directions. *Eur Neuropsychopharmacol* (2016) 26(6):923–37. doi:10.1016/j.euroneuro.2016.01.009
- Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry* (2010) 167:748–51. doi:10.1176/appi.ajp.2010.09091379
- Passingham RE, Rowe JB, Sakai K. Has brain imaging discovered anything new about how the brain works? *Neuroimage* (2013) 66:142–50. doi:10.1016/j.neuroimage.2012.10.079
- Friston K, Daunizeau J, Stephan KE. Model selection and gobbledygook: response to Lohmann et al. *Neuroimage* (2013) 75:275–278; discussion 279–281. doi:10.1016/j.neuroimage.2011.11.064
- Friston KJ. Functional and effective connectivity: a review. *Brain Connect* (2011) 1:13–36. doi:10.1089/brain.2011.0008
- Stephan KE, Friston KJ, Frith CD. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophr Bull* (2009) 35:509–27. doi:10.1093/schbul/sbn176
- Friston KJ. Theoretical neurobiology and schizophrenia. *Br Med Bull* (1996) 52:644–55. doi:10.1093/oxfordjournals.bmb.a011573
- Harrison PJ. The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain* (1999) 122(Pt 4):593–624. doi:10.1093/brain/122.4.593
- Stephan KE, Baldeweg T, Friston KJ. Synaptic plasticity and dysconnection in schizophrenia. *Biol Psychiatry* (2006) 59:929–39. doi:10.1016/j.biopsych.2005.10.005
- Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RA. Hemodynamic signals correlate tightly with synchronized gamma oscillations. *Science* (2005) 309:948–51. doi:10.1126/science.1110948
- Logothetis NK. What we can do and what we cannot do with fMRI. *Nature* (2008) 453:869–78. doi:10.1038/nature06976
- Harel N. Ultra high resolution fMRI at ultra-high field. *Neuroimage* (2012) 62:1024–8. doi:10.1016/j.neuroimage.2012.01.018
- Lowe MJ, Sakaie KE, Beall EB, Calhoun VD, Bridwell DA, Rubinov M, et al. Modern methods for interrogating the human connectome. *J Int Neuropsychol Soc* (2016) 22:105–19. doi:10.1017/S1355617716000060
- Liu Z, Ding L, He B. Integration of EEG/MEG with MRI and fMRI. *IEEE Eng Med Biol Mag* (2006) 25:46–53. doi:10.1109/EMEMB.2006.1657787
- Lang EW, Tome AM, Keck IR, Gorris-Saez JM, Puntinet CG. Brain connectivity analysis: a short survey. *Comput Intell Neurosci* (2012) 2012:412512. doi:10.1155/2012/412512
- Freud S. *Die Traumdeutung*. Vienna: Franz Deuticke (1900).
- Bleuler E. *Dementia Praecox or the Group of Schizophrenias*. New York: International Press (1908).
- APA. *Diagnostic Statistical Manual of Mental Disorders*. Arlington, VA: American Psychiatric Association (2013).
- Sperry RW. Mind-brain interaction: mentalism, yes; dualism, no. *Neuroscience* (1980) 5:195–206. doi:10.1016/0306-4522(80)90098-6
- Northoff G. *Unlocking the Brain: Volume 2: Consciousness*. Oxford: Oxford University Press (2013).
- Kraepelin E. *Dementia Praecox*. New York: Churchill Livingston Inc (1919–1971).
- Bleuler E. *Textbook of Psychiatry*. New York: MacMillan (1924).
- Kosslyn SM, Intriligator JM. Is cognitive neuropsychology plausible? The perils of sitting on a one-legged stool. *J Cogn Neurosci* (1992) 4:96–105. doi:10.1162/jocn.1992.4.1.96
- Price CJ, Friston KJ. Functional ontologies for cognition: the systematic definition of structure and function. *Cogn Neuropsychol* (2005) 22:262–75. doi:10.1080/02643290442000095
- Friston KJ. The disconnection hypothesis. *Schizophr Res* (1998) 30:115–25. doi:10.1016/S0920-9964(97)00140-0
- Andreasen NC. A unitary model of schizophrenia: Bleuler’s “fragmented phrene” as schizencephaly. *Arch Gen Psychiatry* (1999) 56:781–7. doi:10.1001/archpsyc.56.9.781
- Behrens TE, Sporns O. Human connectomics. *Curr Opin Neurobiol* (2012) 22:144–53. doi:10.1016/j.conb.2011.08.005
- Park HJ, Friston K. Structural and functional brain networks: from connections to cognition. *Science* (2013) 342:1238411. doi:10.1126/science.1238411
- Bassett DS, Bullmore E, Verchinski BA, Mattay VS, Weinberger DR, Meyer-Lindenberg A. Hierarchical organization of human cortical networks in health and schizophrenia. *J Neurosci* (2008) 28:9239–48. doi:10.1523/JNEUROSCI.1929-08.2008
- Banyai M, Diwadkar VA, Erdi P. Model-based dynamical analysis of functional disconnection in schizophrenia. *Neuroimage* (2011) 58:870–7. doi:10.1016/j.neuroimage.2011.06.046
- Bressler SL, Tognoli E. Operational principles of neurocognitive networks. *Int J Psychophysiol* (2006) 60:139–48. doi:10.1016/j.ijpsycho.2005.12.008
- Bressler SL. Cortical coordination dynamics and the disorganization syndrome in schizophrenia. *Neuropsychopharmacology* (2003) 28(Suppl 1):S35–9. doi:10.1038/sj.npp.1300145
- Alexander-Bloch A, Lambiotte R, Roberts B, Giedd J, Gogtay N, Bullmore E. The discovery of population differences in network community structure: new methods and applications to brain functional networks in schizophrenia. *Neuroimage* (2012) 59:3889–900. doi:10.1016/j.neuroimage.2011.11.035
- Broca P. Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech). *Bull Soc Anat* (1861) 6:330–57.
- Kertesz A. Wernicke–Geschwind model. *Encyclopedia of Cognitive Science*, New York: John H. Wiley & Sons (2006). doi:10.1002/0470018860.s00307
- Lopez-Munoz F, Boya J, Alamo C. Neuron theory, the cornerstone of neuroscience, on the centenary of the Nobel Prize award to Santiago Ramon y Cajal. *Brain Res Bull* (2006) 70:391–405. doi:10.1016/j.brainresbull.2006.07.010
- Hubel DH, Wiesel TN. Brain mechanisms of vision. *Sci Am* (1979) 241:150–62. doi:10.1038/scientificamerican0979-150
- Chang EF, Bao S, Imaizumi K, Schreiner CE, Merzenich MM. Development of spectral and temporal response selectivity in the auditory cortex. *Proc Natl Acad Sci U S A* (2005) 102:16460–5. doi:10.1073/pnas.0508239102
- Fuster J. Stretch that net or your mind may fall into phrenology. *Cortex* (2006) 42:1048–52. doi:10.1016/S0010-9452(08)70214-4
- Passingham RE, Stephan KE, Kötter R. The anatomical basis of functional localization in the cortex. *Nat Rev Neurosci* (2002) 3:606–16. doi:10.1038/nrn893
- Friston KJ. Models of brain function in neuroimaging. *Annu Rev Psychol* (2005) 56:57–87. doi:10.1146/annurev.psych.56.091103.070311
- Mesulam MM. From sensation to cognition. *Brain* (1998) 121:1013–52. doi:10.1093/brain/121.6.1013

54. Fuster JM. The prefrontal cortex-an update: time is of the essence. *Neuron* (2001) 30:319–33. doi:10.1016/S0896-6273(01)00285-9
55. Stephan KE. On the role of general system theory for functional neuroimaging. *J Anat* (2004) 205:443–70. doi:10.1111/j.0021-8782.2004.00359.x
56. Von Bertalanffy K. *General Systems Theory*. New York: Brazillier (1968).
57. Erdi P. *Complexity Explained*. New York: Springer (2008).
58. Goldman-Rakic PS. Topography of cognition: parallel distributed networks in primate association cortex. *Annu Rev Neurosci* (1988) 11:137–56. doi:10.1146/annurev.ne.11.030188.001033
59. Singer W. Coherence as an organizing principle of cortical functions. *Int Rev Neurobiol* (1994) 37:153–183; discussion 203–157. doi:10.1016/S0074-7742(08)60245-7
60. Nicolelis MA, Lebedev MA. Principles of neural ensemble physiology underlying the operation of brain-machine interfaces. *Nat Rev Neurosci* (2009) 10:530–40. doi:10.1038/nrn2653
61. Singh KD. Which “neural activity” do you mean? fMRI, MEG, oscillations and neurotransmitters. *Neuroimage* (2012) 62:1121–30. doi:10.1016/j.neuroimage.2012.01.028
62. Logothetis NK. The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos Trans R Soc Lond B Biol Sci* (2002) 357:1003–37. doi:10.1098/rstb.2002.1114
63. Logothetis NK, Wandell BA. Interpreting the BOLD signal. *Annu Rev Physiol* (2004) 66:735–69. doi:10.1146/annurev.physiol.66.082602.092845
64. Friston KJ, Li B, Daunizeau J, Stephan KE. Network discovery with DCM. *Neuroimage* (2012) 56:1202–21. doi:10.1016/j.neuroimage.2010.12.039
65. Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, et al. Network modelling methods for FMRI. *Neuroimage* (2011) 54:875–91. doi:10.1016/j.neuroimage.2010.08.063
66. van den Heuvel MP, Hulshoff Pol HE. Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur Neuropsychopharmacol* (2010) 20:519–34. doi:10.1016/j.euroneuro.2010.03.008
67. Stanley ML, Moussa MN, Paolini BM, Lyday RG, Burdette JH, Laurienti PJ. Defining nodes in complex brain networks. *Front Comput Neurosci* (2013) 7:169. doi:10.3389/fncom.2013.00169
68. Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* (1995) 34:537–41. doi:10.1002/mrm.1910340409
69. Lowry EP. The role of time series in econometric model evaluation. In: Kmenta JPRJ, editor. *Evaluation of Econometric Models*. New York: Academic Press (1980). p. 275–307.
70. Sun FT, Miller LM, Desposito M. Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *Neuroimage* (2004) 21:647–58. doi:10.1016/j.neuroimage.2003.09.056S1053811903006062
71. Miller BT, D’Esposito M. Spatial and temporal dynamics of cortical networks engaged in memory encoding and retrieval. *Front Hum Neurosci* (2012) 6:109. doi:10.3389/fnhum.2012.00109
72. Hemmelmann D, Ungureanu M, Hesse W, Wustenberg T, Reichenbach JR, Witte OW, et al. Modelling and analysis of time-variant directed interrelations between brain regions based on BOLD-signals. *Neuroimage* (2009) 45:722–37. doi:10.1016/j.neuroimage.2008.12.065
73. Bressler SL, Seth AK. Wiener-Granger causality: a well established methodology. *Neuroimage* (2011) 58:323–9. doi:10.1016/j.neuroimage.2010.02.059
74. Asemi A, Ramaseshan K, Burgess A, Diwadkar VA, Bressler SL. Dorsal anterior cingulate cortex modulates supplementary motor area in coordinated unimanual motor behavior. *Front Hum Neurosci* (2015) 9:309. doi:10.3389/fnhum.2015.00309
75. Bastos AM, Schoffelen JM. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Front Syst Neurosci* (2015) 9:175. doi:10.3389/fnsys.2015.00175
76. Aersten A, Preisl H. Dynamics of activity and connectivity in physiological neuronal networks. In: Schuster HG, editor. *Non-Linear Dynamics and Neuronal Networks*. New York: VCH Publishing (1991). p. 281–302.
77. Friston K, Moran R, Seth AK. Analysing connectivity with Granger causality and dynamic causal modelling. *Curr Opin Neurobiol* (2013) 23:172–8. doi:10.1016/j.conb.2012.11.010
78. Stephan KE, Penny WD, Moran RJ, Den Ouden HE, Daunizeau J, Friston KJ. Ten simple rules for dynamic causal modeling. *Neuroimage* (2010) 49:3099–109. doi:10.1016/j.neuroimage.2009.11.015
79. Stephan KE, Harrison LM, Kiebel SJ, David O, Penny WD, Friston KJ. Dynamic causal models of neural system dynamics: current state and future extensions. *J Biosci* (2007) 32:129–44. doi:10.1007/s12038-007-0012-5
80. Sun FT, Miller LM, Desposito M. Measuring temporal dynamics of functional networks using phase spectrum of fMRI data. *Neuroimage* (2005) 28:227–37. doi:10.1016/j.neuroimage.2005.05.043
81. Lee U, Lee H, Müller M, Noh G-J, Mashour GA. Genuine and spurious phase synchronization strengths during consciousness and general anesthesia. *PLoS One* (2012) 7:e46313. doi:10.1371/journal.pone.0046313
82. Schreiber T, Schmitz A. Surrogate time series. *Phys D Nonlinear Phenom* (2000) 142:346–82. doi:10.1016/S0167-2789(00)00043-9
83. Pal D, Hambrecht-Wiedbusch V, Silverstein B, Mashour G. Electroencephalographic coherence and cortical acetylcholine during ketamine-induced unconsciousness. *Br J Anaesth* (2015) 114:979–89. doi:10.1093/bja/aeu095
84. Uhlhaas PJ, Roux F, Rodriguez E, Rotarska-Jagiela A, Singer W. Neural synchrony and the development of cortical networks. *Trends Cogn Sci* (2010) 14:72–80. doi:10.1016/j.tics.2009.12.002
85. Diba K, Amarasingham A, Mizuseki K, Buzsaki G. Millisecond timescale synchrony among hippocampal neurons. *J Neurosci* (2014) 34:14984–94. doi:10.1523/JNEUROSCI.1091-14.2014
86. Buxton RB. Dynamic models of BOLD contrast. *Neuroimage* (2012) 62:953–61. doi:10.1016/j.neuroimage.2012.01.012
87. Lee U, Ku S, Noh G, Baek S, Choi B, Mashour GA. Disruption of frontal-parietal communication by ketamine, propofol, and sevoflurane. *Anesthesiology* (2013) 118:1264–75. doi:10.1097/ALN.0b013e31829103f5
88. Moon J-Y, Lee U, Blain-Moraes S, Mashour GA. General relationship of global topology, local dynamics, and directionality in large-scale brain networks. *PLoS Comput Biol* (2015) 11:e1004225. doi:10.1371/journal.pcbi.1004225
89. Paus T. Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat Rev Neurosci* (2001) 2:417–24. doi:10.1038/35077500
90. Vogt BA. Midcingulate cortex: structure, connections, homologies, functions and diseases. *J Chem Neuroanat* (2016) 74:28–46. doi:10.1016/j.jchemneu.2016.01.010
91. Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M. Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *J Neurosci* (2008) 28:10056–61. doi:10.1523/JNEUROSCI.1776-08.2008
92. Carter CS, Macdonald AW III, Ross LL, Stenger VA. Anterior cingulate cortex activity and impaired self-monitoring of performance in patients with schizophrenia: an event-related fMRI study. *Am J Psychiatry* (2001) 158:1423–8. doi:10.1176/appi.ajp.158.9.1423
93. Bakshi N, Pruiett P, Radwan J, Keshavan M, Rajan U, Zajac-Benitez C, et al. Inefficiently increased anterior cingulate modulation of cortical systems during working memory in young offspring of schizophrenia patients. *J Psychiatr Res* (2011) 45:1067–76. doi:10.1016/j.jpsychires.2011.01.002
94. Woodcock EA, Wadehra S, Diwadkar VA. Network profiles of the dorsal anterior cingulate and dorsal prefrontal cortex in schizophrenia during hippocampal-based associative memory. *Front Syst Neurosci* (2016) 10:32. doi:10.3389/fnsys.2016.00032
95. Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* (1997) 6:218–29. doi:10.1006/nimg.1997.0291
96. O’Reilly JX, Woolrich MW, Behrens TE, Smith SM, Johansen-Berg H. Tools of the trade: psychophysiological interactions and functional connectivity. *Soc Cogn Affect Neurosci* (2012) 7:604–9. doi:10.1093/scan/nss055
97. Horwitz B, Warner B, Fitzer J, Tagamets MA, Husain FT, Long TW. Investigating the neural basis for functional and effective connectivity. Application to fMRI. *Philos Trans R Soc Lond B Biol Sci* (2005) 360:1093–108. doi:10.1098/rstb.2005.1647
98. Wadehra S, Pruiett P, Murphy ER, Diwadkar VA. Network dysfunction during associative learning in schizophrenia: increased activation, but decreased connectivity: an fMRI study. *Schizophr Res* (2013) 148:38–49. doi:10.1016/j.schres.2013.05.010
99. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* (1969) 37:424–38. doi:10.2307/1912791
100. David O, Guillemain I, Saitlet S, Rey S, Deransart C, Segebarth C, et al. Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS Biol* (2008) 6:2683–97. doi:10.1371/journal.pbio.0060315

101. Friston K. Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biol* (2009) 7:e33. doi:10.1371/journal.pbio.1000033
102. Borjigin J, Lee U, Liu T, Pal D, Huff S, Klarr D, et al. Surge of neurophysiological coherence and connectivity in the dying brain. *Proc Natl Acad Sci U S A* (2013) 110:14432–7. doi:10.1073/pnas.1308285110
103. Seth AK, Chorley P, Barnett LC. Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. *Neuroimage* (2013) 65:540–55. doi:10.1016/j.neuroimage.2012.09.049
104. Tang W, Bressler SL, Sylvester CM, Shulman GL, Corbetta M. Measuring Granger causality between cortical regions from voxelwise fMRI BOLD signals with LASSO. *PLoS Comput Biol* (2012) 8:e1002513. doi:10.1371/journal.pcbi.1002513
105. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. *Neuroimage* (2003) 19:1273–302. doi:10.1016/S1053-8119(03)00202-7
106. Daunizeau J, David O, Stephan KE. Dynamic causal modelling: a critical review of the biophysical and statistical foundations. *Neuroimage* (2011) 58:312–22. doi:10.1016/j.neuroimage.2009.11.062
107. Frassle S, Paulus FM, Krach S, Jansen A. Test-retest reliability of effective connectivity in the face perception network. *Hum Brain Mapp* (2016) 37:730–44. doi:10.1002/hbm.23061
108. Diwadkar VA, Bakshi N, Gupta G, Pruitt P, White R, Eickhoff SB. Dysfunction and dysconnection in cortical-striatal networks during sustained attention: genetic risk for schizophrenia or bipolar disorder and its impact on brain network function. *Front Psychiatry* (2014) 5:50. doi:10.3389/fpsy.2014.00050
109. Jagtap P, Diwadkar VA. Effective connectivity of ascending and descending frontothalamic pathways during sustained attention: complex brain network interactions in adolescence. *Hum Brain Mapp* (2016) 37(7):2557–70. doi:10.1002/hbm.23196
110. Stephan KE, Tittgemeyer M, Knosche TR, Moran RJ, Friston KJ. Tractography-based priors for dynamic causal models. *Neuroimage* (2009) 47:1628–38. doi:10.1016/j.neuroimage.2009.05.096
111. Seghier ML, Zeidman P, Neufeld NH, Leff AP, Price CJ. Identifying abnormal connectivity in patients using dynamic causal modeling of FMRI responses. *Front Syst Neurosci* (2010) 4:142. doi:10.3389/fnsys.2010.00142
112. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* (2003) 160:636–45. doi:10.1176/appi.ajp.160.4.636
113. Kaminski M, Ding M, Truccolo WA, Bressler SL. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern* (2001) 85:145–57. doi:10.1007/s004220000235
114. Diwadkar VA. Critical perspectives on causality and inference in brain networks: allusions, illusions, solutions? Comment on: “Foundational perspectives on causality in large-scale brain networks” by M. Mannino and S.L. Bressler. *Phys Life Rev* (2015) 15:141–4. doi:10.1016/j.plrev.2015.10.018
115. Mannino M, Bressler SL. Foundational perspectives on causality in large-scale brain networks. *Phys Life Rev* (2015) 15:107–23. doi:10.1016/j.plrev.2015.09.002

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Silverstein, Bressler and Diwadkar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.